

Il modello Ready for commons nella pratica della ricerca

Andrea Raimondi, *OSD opensensorsdata*, andrea@osd.tools

Luca Corsato, *OSD opensensorsdata*, luca@osd.tools

Simone Cortesi, *OSD opensensorsdata*, simone@cortesi.com

OSD opensensorsdata¹ ha di recente avviato uno studio di modelli microeconomici per lo sviluppo delle pratiche di condivisione delle risorse informative. Obiettivo del presente position paper è quello di introdurre il modello Ready for Commons e mostrare come questo modello sia in grado di far fronte ad alcune difficoltà dei modelli di riuso tradizionali, senza che ciò vada a scapito della missione di ricerca accademica e delle sue pratiche contemporanee.

1. La ricerca come Riuso

L'attività di un ricercatore si concentra nella pubblicazione di articoli scientifici che costituiscono la letteratura fondamentale per l'avanzamento della conoscenza. Nel processo di produzione della conoscenza gli articoli sono l'unità minima di valore: costituiscono l'elemento di valutazione della qualità di produzione e sono ciò in base a cui i processi di produzione riescono ad attrarre fondi. Tale unità minima non è isolata da altri processi produttivi. L'attendibilità di un articolo, infatti, è determinata anche dalla bibliografia cioè dalla capacità di riusare risorse presenti in letteratura, di inserirsi in un contesto di dibattito, e di promuoverlo attraverso contributi informativi nuovi.

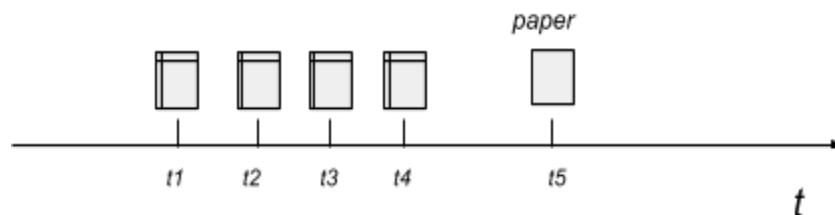
¹ OSD opensensorsdata srl <https://osd.tools>

2. Ontologia del Riuso: una visione tradizionale

In ambito accademico, il sistema di riuso della conoscenza identifica il valore di produzione solo nello stato finale del processo, i.e. l'articolo. L'assunto di questo ragionamento è che gli elementi minimi di costruzione dell'articolo non costituiscano di per sé conoscenza, né abbiano di per sé valore.

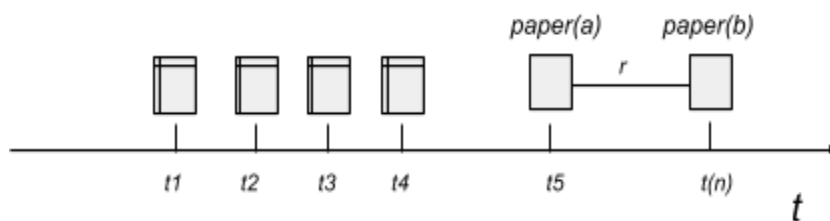
Il processo, e i suoi step temporali, sono considerati come elementi di valore solo in relazione allo step finale: servono a creare un oggetto minimo certificato, condivisibile e riusabile. Questo è il loro uso e nessun riuso di questi elementi derivati è possibile, poiché l'articolo, e non ciò di cui è costituito, rappresenta conoscenza.

Data tale ontologia, la conoscenza è la proprietà di un oggetto non esteso nel tempo. L'indice di riuso è determinato dalle citazioni, il numero di volte in cui un oggetto diventa parte di un'altra unità minima di valore. L'ontologia degli oggetti di conoscenza che abbiamo appena presentato può essere rappresentata in questo modo.



L'articolo, come unità minima, esiste al tempo t_5 , ma non esiste prima. Ciò che esiste prima, dal tempo t_1 a t_4 , è ogni step di ricerca che conduce al risultato finale. Esistendo solo al tempo t_5 , l'articolo

non ha parti temporali. È un oggetto tridimensionale.² Il riuso, di conseguenza, è una relazione r che interessa solo l'oggetto al tempo t_5 e che mette in relazione questo con un altro articolo al tempo $t(n)$.



Dal momento che l'articolo è considerato come un oggetto tridimensionale, la relazione di riuso non coinvolge nessuno degli step, né ciò che in questi step viene prodotto. Questa ontologia, tuttavia, presenta alcuni limiti. Vediamo quali.

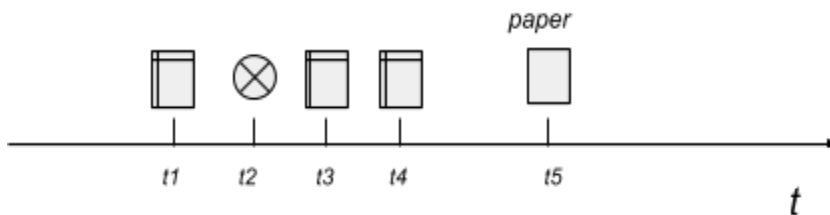
3. Riuso e Negatività

Se l'unità minima di valore è il prodotto finale del processo di ricerca, gli step precedenti non aggiungono alcun valore a quello che il prodotto finale già presenta. Sono, per così dire, epifenomeni. Non c'è nulla di informativo che non sia già dato nel prodotto finale. Ci sono alcuni casi, però, in cui questa conclusione appare ingiustificata. Tali sono i casi di negatività.

I casi di negatività si dividono in due categorie: dati negativi, e articoli negativi. Vengono definiti *dati negativi*, quei dati che mostrano risultati in cui l'effetto aspettato non si mostra, che mostrano risultati controversi o neutri rispetto all'ipotesi da provare, e che, a causa di ciò, non sono mai parte (concreta) della pubblicazione finale (Andow 2003; *inter alia*).

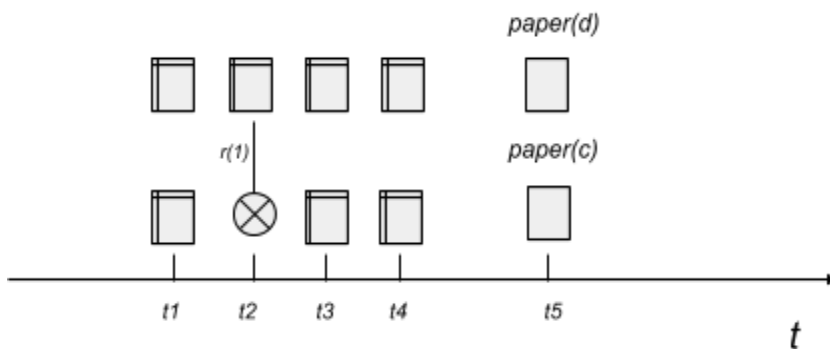
² Sulla dottrina del tridimensionalismo e sulla persistenza degli oggetti tridimensionale si veda (Ney 2014:170-189).

Supponiamo, ad esempio, di voler testare l'effetto di un additivo per il suolo in relazione al rating di crescita di un tipo di pianta. Se l'esperimento non mostra variazioni significative, il dataset creato al tempo t_2 a seguito della sperimentazione finisce inutilizzato negli archivi di laboratorio.

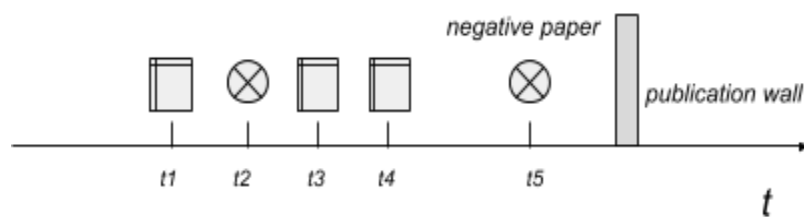


Lo step a t_2 è un caso di dato negativo; esiste come step di produzione, ma non fa parte della conoscenza presente nello step finale. Per l'approccio ontologico sopra descritto, dunque, questi dati non codificano nessuna informazione rilevante.

Questa conclusione è tuttavia problematica. I dati negativi sono informativi perché permettono ad altri processi produttivi di conoscere quali metodologie non sono adeguate al raggiungimento di certe ipotesi di ricerca. Questo permette di diminuire i costi di produzione della conoscenza riusando elementi di produzione della conoscenza creati all'interno di altri processi.



I dati negativi sono informativi in un altro senso, poiché permettono di correggere falsi-positivi molto più velocemente. Permettono, in particolare, il riuso in contesti di *articoli negativi*, cioè articoli il cui obiettivo è quello di riprodurre risultati sperimentali. Questi, tuttavia, raggiungono molto raramente la fase di pubblicazione. Dati i costi di produzione del prodotto conoscitivo –l’articolo– la maggior parte degli istituti di ricerca predilige l’investimento di risorse in articoli positivi, che restituiscono un maggiore guadagno in termini di citazione e prestigio (Fanelli 2012; *inter alia*).



Se adottiamo il modello precedentemente descritto nulla di questi dati accresce il valore conoscitivo ottenuto dal processo produttivo. In quel modello il valore conoscitivo è, in modo standard, inteso come citabilità del prodotto finale. Il suo valore cresce al crescere delle citazioni, ma non viene mai esteso. Definiamo come *estensione di valore* il numero di volte in cui una parte dell’oggetto, non l’oggetto stesso, diventa parte di un’altra unità.

Se l’articolo è l’unità minima di valore di conoscenza, nessuna estensione di valore è possibile. Questo perché tale unità non ha parti. Dunque nulla della capacità informativa di riuso dei dati negativi entra nel valore del prodotto finale del processo.

In secondo luogo, la relazione r di riuso viene intesa, in quel modello, come relazione tra unità minime, i.e. gli articoli. Non avendo tali unità parti che corrispondono agli stage produttivi, nessuna relazione r di riuso può essere ammessa nel modello.

In ultima istanza, il modello non offre nessuna soluzione al problema degli articoli negativi. Se, infatti, il valore produttivo è pensato come proprietà del prodotto finale, se il prodotto finale ha valore negativo, nessuna unità minima può in principio raggiungere lo step finale di pubblicazione. E nessuna soluzione è presente che consenta di valorizzare le parti di produzione, perché queste, come già detto, non hanno valore se non nello step finale di pubblicazione.

4. Ready for commons

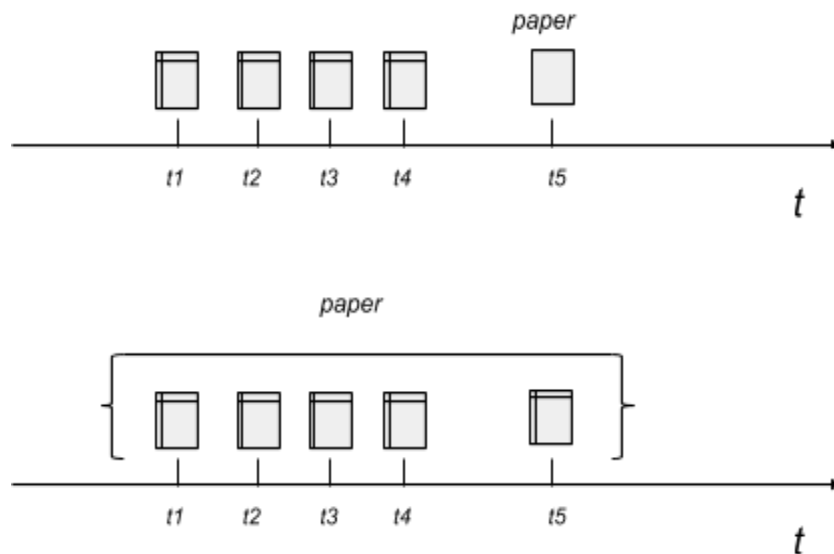
L'obiettivo di questa sezione è mostrare come il modello Ready for commons (RFC) sia in grado di superare le limitazioni connesse al modello precedente. Il modello RFC ha come obiettivo quello di rendere tutti i risultati, che accadono negli step precedenti alla pubblicazioni, disponibili al riuso per altri processi di pubblicazione. L'idea alla base del modello RFC è quella di rendere ogni step del processo *ready for* relazioni di condivisione, in modo da incrementare la pubblicazione delle risorse e abbassare i costi di produzione. Iniziamo, dunque, da come RFC comprende l'oggetto di conoscenza.

4.1 Parti di Processi

Il modello precedente comprendeva gli oggetti informativi come oggetti tridimensionali: l'articolo è solamente quell'oggetto al tempo t_5 , diverso da ogni step di produzione. Secondo RFC,

invece, l'articolo è l'intero processo di produzione. Poiché il processo di produzione è esteso nel tempo, RFC sostiene una concezione quadridimensionalista degli oggetti di conoscenza.³

Secondo questo approccio, l'oggetto al tempo t_5 è una *parte temporale* dell'oggetto di conoscenza, e dunque non identico con esso. Gli step che portano alla parte temporale in t_5 sono anch'essi parti temporali. Insieme, costituiscono l'oggetto di conoscenza la cui natura non può essere slegata né dalla dimensione spaziale né da quella temporale. Per i quadridimensionalisti, infatti, il tempo funziona in modo analogo allo spazio, è una dimensione. Spostarsi nel tempo è come visitare luoghi fisici collegati tra di loro da relazioni spazio--temporali. Comparete:

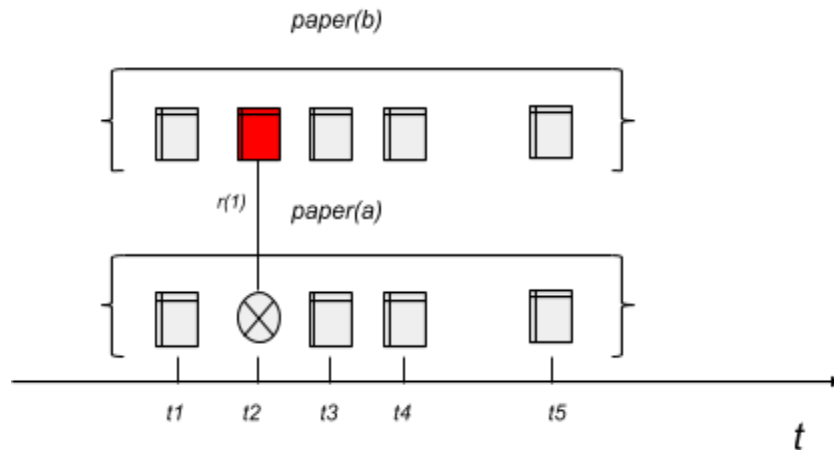


4.2 Riuso e Produzione: una visione unitaria

Vediamo come utilizzare questo modello per affrontare i problemi esposti in (3.). In primo luogo, dal momento che l'articolo ha sia parti fisiche che parti temporali, le sue parti temporali possono

³ Sulla dottrina del quadridimensionalismo si veda ancora (Ney 2014:170-189).

essere riusate tanto quanto sono riusate le parti fisiche del articolo nel modello precedente. Dal momento che ciò che viene riusato non è solo l'articolo nella sua interezza, ma anche parti temporali di produzione di esso, le relazioni r di riuso possono essere istanziate da molti più oggetti, i.e. in questo caso ogni, o alcune, parti temporali.



Una prima e importante considerazione è che il modello RFC, rendendo *ready* alla condivisione ogni prodotto ad ogni segmento temporale, non comprende la condivisione come posteriore al processo di produzione. Non siamo più all'interno di un framework in cui prima si produce e poi si condivide. Al contrario, la condivisione è condivisione di parti produttive.

Nel modello RFC, due processi produttivi distinti, come nel caso sopra, *letteralmente* condividono due parti –quelle collegate da $r(i)$ nel caso sopra descritto. La condivisione, da processo diffusivo della conoscenza, diventa processo costitutivo della produzione di conoscenza. Vediamo come questo modello si applica al problema dei dati negativi.

La parte temporale t_2 dell'articolo(a) include la produzione di un dato negativo. Nel modello precedente il potenziale informativo di quello step non poteva essere utilizzato perché nessuna relazione r di riuso poteva essere istanziata da uno step del processo.

In questo caso, invece, dal momento che l'articolo è esteso nel tempo e ha parti temporali, la parte temporale a t_2 può istanziare una relazione $r(t)$ di riuso. Dunque, il modello RFC permette l'uso di dati negativi che altrimenti non sarebbero valorizzati.

Cosa dire degli articoli negativi? se lo scopo dei articoli negativi è quello di mostrare “unsuccessful solutions to the problem” (Heumann 1956:36), un archivio di dataset negativi può svolgere lo stesso scopo senza dover vincolare il processo produttivo ad una pubblicazione negativa –evitando di conseguenza il problema dei finanziamenti ai prodotti di validazione di risultati già esistenti ma da confermare. In tale modello la validazione dei risultati e il controllo sui falsi positivi avviene in fase di produzione sulla base dei dati negativi condivisi da altri processi di produzione.

Abbiamo, dunque, un'ontologia più parsimoniosa; dobbiamo includere dati negativi nella nostra ontologia, dati i fatti, ma possiamo ridurre gli articoli negativi ai dati negativi, riducendo il numero di entità che popolano la nostra ontologia. Se è possibile ridurre *tutti* gli articoli negativi o solo *alcuni* è una domanda che richiede una valutazione caso per caso e eccede lo scopo di questo articolo.

4.3 Limiti del modello

Nella sezione precedente abbiamo visto come il modello RFC possa offrire una soluzione al problema dei dati negativi. Ne permetterebbe il riuso e consentirebbe alla comunità scientifica di usufruire di informazioni cruciali per modellare altri processi produttivi, aumentando la capacità di raffinamento delle ipotesi di ricerca e abbassando i costi di riproduzione dei risultati.

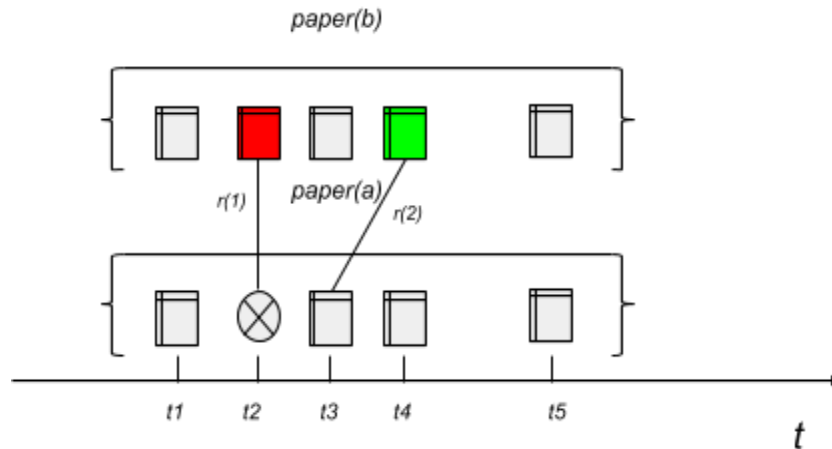
Il modello RFC, nonostante non sia costruito al fine di risolvere il problema dei dati negativi è, in questo caso, un passo nella direzione di valorizzare tutti i dati, cioè tutte le parti del processo, come se avessero lo stesso valore delle pubblicazione. La tesi del modello RFC, infatti, è che tutte le parti del processo possono essere “messe a valore”. In questa sezione analizziamo alcuni limiti del modello.

La prima domanda è su come valorizzare tutte le parti del processo. Se l'articolo è l'intero oggetto quadrimensionale, sembra che il valore di riuso di ogni parte non sia diverso dal valore di riuso dell'intero oggetto. L'idea del modello RFC consente semplicemente di parcellizzare il valore del processo in valori minimali delle parti. Esaminiamo quindi questa proposta.

Nel primo modello descritto il riuso era una relazione tra prodotti pubblicati. In questo caso il riuso riguarda le parti del prodotto. Ready for commons propone a chiunque svolga attività di ricerca di creare un *data for paper*.

Data for paper è l'insieme di un pacchetto dati e la loro descrizione metodologica e di riuso, prodotti in una parte temporale del articolo, da sottoporre ad un comitato scientifico. I data for paper sono validati e pubblicati su una rivista che, assieme alla descrizione associata, entreranno nella letteratura scientifica a supporto dei articolo: data for paper, appunto. Può il riuso delle parti seguire la stessa nozione di riuso del prodotto finale? La risposta è no.

Se la parte in questione è funzionale alla pubblicazione dell'articolo (ipotesi A), allora il riuso è incoerente rispetto alla funzionalità. Se invece la parte in questione non è funzionale (ipotesi B, vedi dati negativi) allora la pubblicazione di tale parte, nonostante permetta l'abbassamento dei costi di produzione di altri processi, aumenta il costo di produzione nel processo di partenza. Vediamo con attenzione perché .



Nel caso sopra modellato abbiamo un caso di riuso di parte di processo: la parte a t_3 è in relazione di condivisione con la parte t_4 dell'articolo(b). Lo scopo di RFC è quello di dare valore a fasi di produzione della conoscenza che non siano la sola fase finale, i.e. pubblicazione. In questo caso, poter valorizzare un set di dati –che si userà in seguito per un articolo– pubblicandolo già da prima in una repository. Secondo la proposta in esame, t_3 viene trattato come un prodotto finale. Viene cioè curato nei suoi aspetti metodologici e proposto come oggetto di pubblicazione. Il problema è il seguente.

La repository o attribuisce DOI o non lo attribuisce. Se lo attribuisce, il set di dati con attribuzione può essere pubblicato, ma non può poi essere usato per pubblicare successivamente l'articolo per il quale era finalizzato. Questo perché i dati pubblicati in sede finale non devono essere pubblicati precedentemente da altre parti.

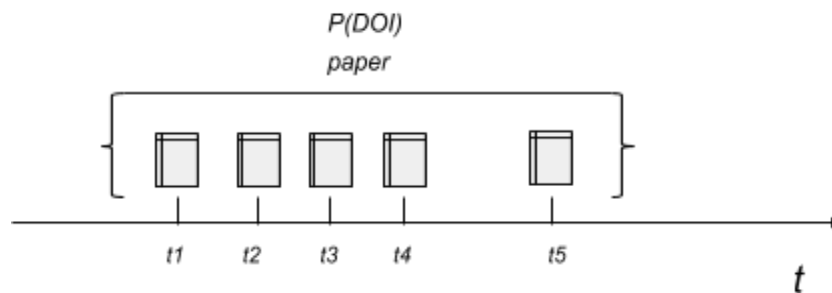
Se invece non lo attribuisce, quei dati senza attribuzione possono essere riutilizzati per la pubblicazione finale non solo dal soggetto che li ha prodotti, ma anche da altri. E dunque è possibile

che il soggetto non possa pubblicare al tempo t_5 un articolo in fase finale, perché altri soggetti, possibilmente, hanno già incluso quei dati in una loro pubblicazione al tempo $((t_5)-1)$.⁴

Per riassumere, il modello RFC vuole mettere a valore ogni singola parte temporale, ma tratta il valore prodotto nella forma in modo isomorfo al modello descritto nella sezione 3. In questo modo, però, include alcuni limiti tecnici collegati ai processi di pubblicazione e citazione caratteristici del modello in sezione 3 senza offrire soluzione alcuna.

5. Parti e Controparti

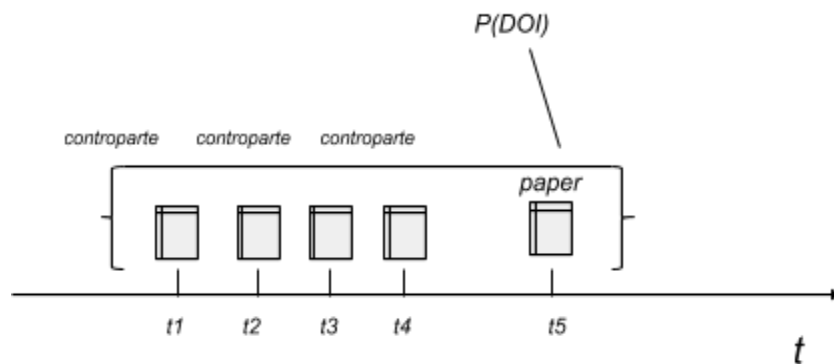
Il modello RFC presentato sopra concettualizza l'articolo come un'entità quadrimensionale. L'articolo è la somma delle sue parti temporali. Ed ogni proprietà dell'articolo è una proprietà che nessuna delle sua parti ha. Il DOI di pubblicazione, ad esempio, può essere istanziato solo dall'entità estesa.



⁴ Un altro problema è presente. Valorizzare ogni segmento produttivo ha costi di tempo e soldi che, se usati per tale attività, vengono a mancare per attività successive, cioè la produzione di altre parti temporali. Questo è un importante limite tecnico. Il cuore del problema è il seguente: la condivisione necessita di un modello di costo altrettanto condiviso. In mancanza di tale modello, la condivisione diventa onerosa al punto di minare lo scopo di produzione della parte temporale, cioè quella di costituzione del processo. Non ci occuperemo, tuttavia, di analizzare questo problema nel corso di questo articolo.

Il problema dell'ipotesi A è ora chiaro. La parte temporale in t_3 non può, per questo modello, avere un DOI, poiché questo è istanziato solo come proprietà dell'intero processo esteso. Se vogliamo modellare una relazione r di riuso tra parti temporali, dunque, questa non può essere una relazione di pubblicazione.

Il modello ha anche una componente controintuitiva. Non vogliamo dire, infatti, che l'articolo è già presente al momento in cui ci si inizia a lavorare. Vogliamo dire che il articolo è un prodotto finale. Ciò che vogliamo evitare è solamente l'assunto che il valore della conoscenza risieda *solo* nel prodotto finale. Possiamo modificare il modello in questo modo.



Chiamiamo questo modello RFC*. Un soluzione al problema del modello RFC è la seguente. L'articolo non è l'intero processo esteso. Il processo esteso è il processo di produzione. In questo modo, processo di produzione e prodotto non sono qualitativamente identici. Questo permette di riconoscere parti temporali di riuso, ma di mantenere l'intuizione che l'articolo possa istanziare proprietà che le parti non istanziano, e viceversa.

Le parti temporali in questo modello sono controparti temporali dell'articolo, stadi di produzione simili, in maniera rilevante, e collegati causalmente allo stadio finale che chiamiamo articolo.

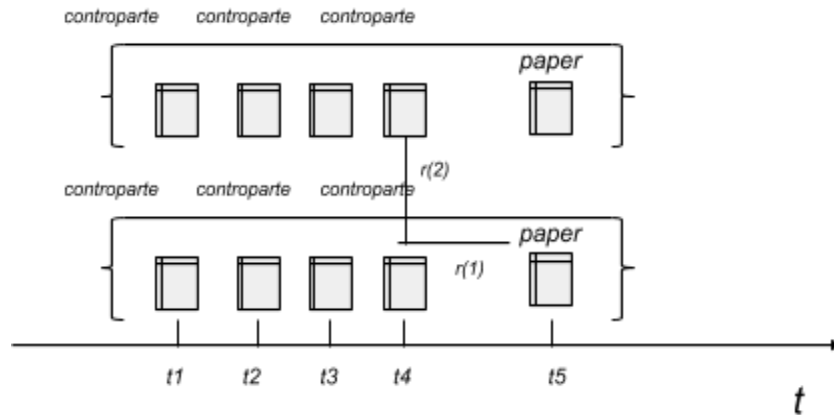
5.1 Controparti ed Estensione di Valore

Questo approccio permette una modellazione più consistente delle parti di riuso. Se le parti sono controparti non è necessario che tutte le parti abbiano proprietà uguali, poiché la relazione di controparte non implica identità stretta ma similarità. È dunque possibile che la controparte in t_5 istanzi la proprietà di avere un DOI mentre le altre controparti no. Non è dunque più necessario trattare il valore di riuso delle parti in modo isomorfo al valore di riuso della parte finale come pensato nell'ipotesi(A).

In secondo luogo, una modellazione più raffinata delle relazione di riuso. Nei casi precedentemente trattati, il riuso di una parte era la condivisione di quella parte del processo con un altro processo. Se, invece, le parti sono controparti temporali nel processo, la relazione di riuso può essere istanziata tra le parti di uno stesso processo. Dunque, anche tra le controparti dell'articolo e l'articolo stesso. Come questa modellazione può aiutare a risolvere il problema dell'ipotesi(A)?

Il problema dell'ipotesi(A) era quello di coniugare la condivisione con le altre attività del processo di produzione. Il problema, in particolare, era dovuto al fatto che ogni parte temporale è parte di un oggetto più grande. Ogni parte era quindi pensata e finalizzata alla costruzione dell'articolo; era infatti parte costitutiva. Il valore positivo di RFC era la sua capacità di sfruttare il valore di alcune parti temporali non finalizzate alla costituzione, come i dati negativi. Il modello, però,

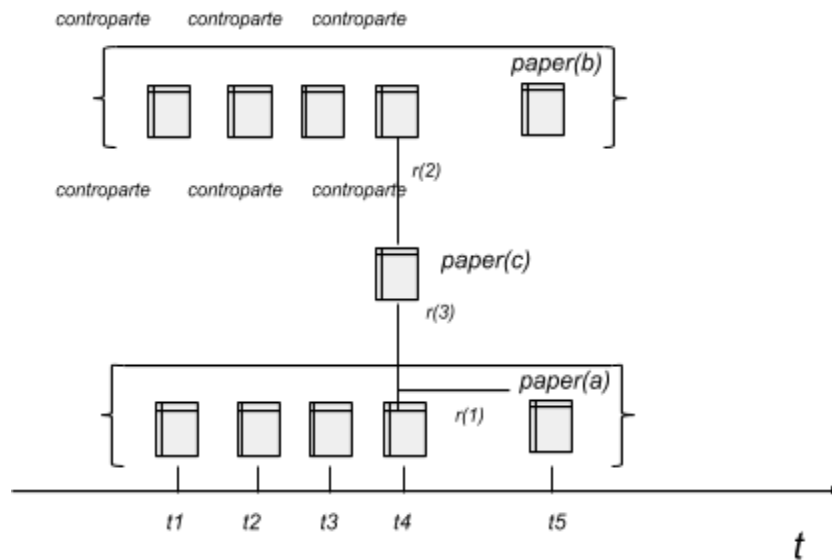
non faceva altro che valorizzare by-product del processo. Il modello a controparti, invece, è profondamente diverso.



In questo modello il riuso non è una relazione tesa a massimizzare anche gli elementi di scarto del processo, ma a massimizzare i tipi di processi produttivi a cui ogni controparte può partecipare. Una controparte, t_4 ad esempio, può essere riusata sia per la produzione di un articolo all'interno del processo di produzione originario, sia per la produzione di un articolo relativo ad un diverso processo di produzione. Vediamo ora quale soluzione questo modello adotta per il problema relativo all'ipotesi(A), e se questa sia soddisfacente.

5.2 Estensione ed il Problema della Pubblicazione

Il punto centrale dell'ipotesi(A) è che la parte del processo è funzionale alla pubblicazione dell'articolo. Il problema è che tale parte, se pubblicata prima, previene una futura pubblicazione a partire da quegli stessi dati. Poiché ogni controparte di riuso non può essere pubblicata, per i problemi esposti in 4.3, la rappresentazione va modificata come segue:



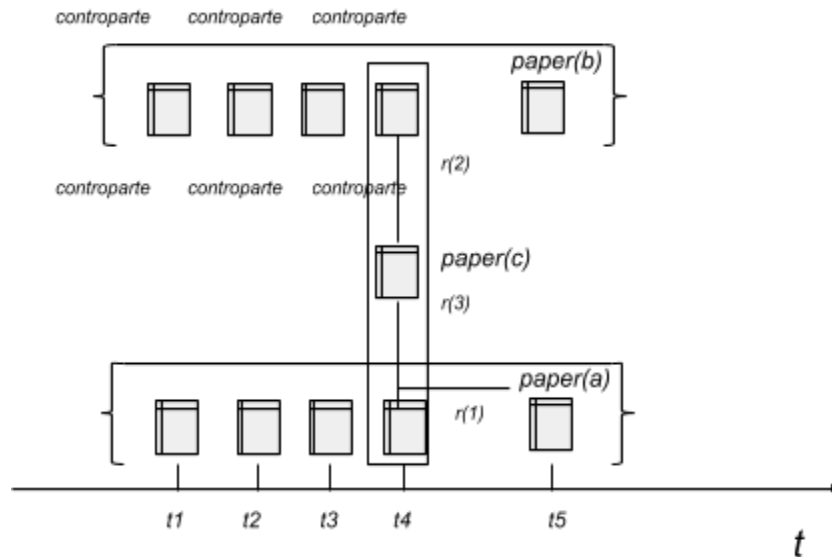
Facciamo un esempio. Supponiamo che l'articolo(a) debba fornire informazioni riguardo l'effetto di diversi additivi sulla crescita della pianta X, mentre l'articolo(b) debba fornire informazioni riguardo l'effetto di un additivo 4 sulla pianta X in diversi contesti ecologici.

L'articolo(a) ha come controparti $t_1(a)..t_4(a)$. Per ogni controparte, lo studio esamina l'effetto sulla pianta X degli additivi 1,2,3 e 4, in modo tale che $t_1(a)=[1,X]$, $t_2(a)=[2,X]$ e così via. L'articolo(b) ha come controparti $t_1(b)..t_4(b)$. Per ogni controparte, lo studio esamina l'effetto dell'additivo 4 su X in contesti J,Y,Z e W, in modo tale che $t_1(b)=[4,X(j)]$, $t_2(b)=[4,X(y)]$ e così via. Osserviamo ora le controparti $t_4(a)$ e $t_4(b)$. La prima corrisponde alla tupla $[4,X]$ mentre la seconda alla tupla $[4,X(w)]$. Mentre la prima esamina l'effetto dell'additivo 4, tra diversi additivi, sulla pianta X, la seconda esamina l'effetto dell'additivo 4 sulla pianta X in diversi contesti. Il riuso delle controparti permette di generare una domanda di ricerca non riducibile né all'articolo(a) né all'articolo(b), cioè, ad esempio,

come la pianta X processa lo stesso additivo in casi di contesto controllato e in caso di variazione del contesto. Chiamiamo questo articolo(c).

È importante notare che l'articolo(a) e l'articolo(c) hanno una controparte in comune, ma sono in ogni caso articolo differenti. Ciò significa che il processo di produzione non genera necessariamente un articolo dato un set di controparti, ma per ogni set di controparti queste possono generare diversi articoli. La stessa controparte può stare in relazione di riuso sia con l'articolo(a) che con l'articolo(b) senza per questo non rientrare fuori dai confini della attività di produzione programmate.

Così modellata, la relazione di riuso non implica che sia la stessa controparte ad avere un DOI. Non abbiamo più, dunque, il problema esposto in (4.3). Al contrario, il DOI rimane proprietà dell'articolo –il (c) in questo caso. Non abbiamo nemmeno più il secondo problema esposto in (3) relativo ai costi di produzione. La produzione dell'articolo(c), infatti, avviene in parallelo e in condivisione tra i due processi di produzione che ora hanno due controparti in comune e che generano un prodotto finale (c) diverso da (a) e (b).



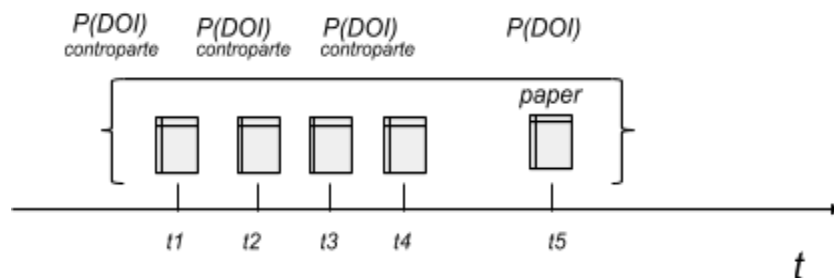
Non sono, in questo modello, i dati ad essere possibili by-product del processo di produzione, ma è l'articolo ad esserlo. Il valore non è qui nello stadio finale e nelle sue citazioni, ma in quanta *estensione di valore* delle attività di produzione ogni parte è capace di trasmettere. Di conseguenza, il valore di investimento delle attività di produzione non è più calcolato in termini di capacità di pubblicazione, ma in termini di capacità di produzione di risorse. Queste, non il articolo, sono le unità minime che, nel modello RFC*, non sono isolate da altri processi produttivi, ma al contrario li costituiscono.

Questo modello, tuttavia, non è esente da difficoltà. Infatti, richiede un ripensamento di quella che viene considerata attività di ricerca. Il costo di adozione del modello è decisamente oneroso. Implica un ripensamento dei modelli di produzione della conoscenza. In particolare, uno spostamento di attenzione dalla produzione di pubblicazioni alla produzione di dati e metodologie, e un conseguente spostamento di valutazione della produzione dalla diffusione di valore all'estensione di valore.

Quanto questa soluzione sia praticabile è un tema troppo approfondito ed esteso da essere trattato in questa sede. Rimane il punto, tuttavia, di una soluzione teoricamente consistente ma praticamente poco implementabile. Infatti è una soluzione che, richiede di abbandonare la missione primaria, considerata quella di pubblicare. È dunque una soluzione promettente, ma che non risolve il problema dell'ipotesi(A) e che, in più, non rispetta i criteri di identità dell'ambito in cui viene implementata. Vediamo ora un'altra soluzione.

5.3 Una Soluzione Semplice

L'ipotesi(A) si basa sull'assunto che vi sia sempre una stretta correlazione tra produzione di conoscenza e pubblicazione. Ci sono, d'altra parte, casi in cui tale correlazione viene a mancare. Pensiamo a quei contesti di ricerca i cui finanziamenti non sono diretti ad attività di pubblicazione, ma di pura produzione di alcuni oggetti di conoscenza, e.g. dati. Il comparto della ricerca privata è uno di questi casi.



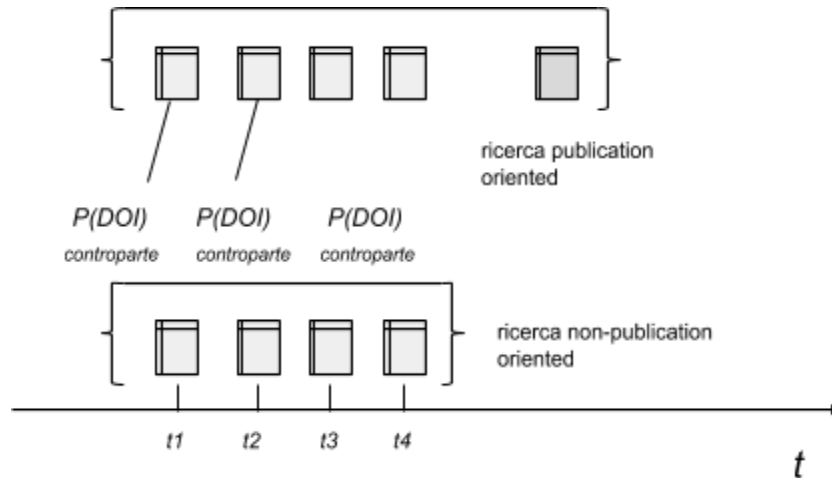
Dal momento che ogni controparte è ontologicamente equivalente ad ogni altra, non c'è motivo di supporre che la pubblicazione abbia un qualche tipo di priorità produttiva sulla controparti. Per

risolvere il problema dell'ipotesi(A), quindi, è necessario solamente estendere l'attribuzione della proprietà DOI a tutte, o alcune, delle controparti.

Supponiamo ad esempio di avere un comparto di ricerca privato i cui prodotti non sono, per stipulazione, prodotti finalizzati alla pubblicazione. Ciò non significa che da tali prodotti non possa esserci una pubblicazione; solamente che le attività permesse all'interno di tale comparto non sono mirate alla pubblicazione. Come estendere il valore informativo degli oggetti di conoscenza in questo caso?

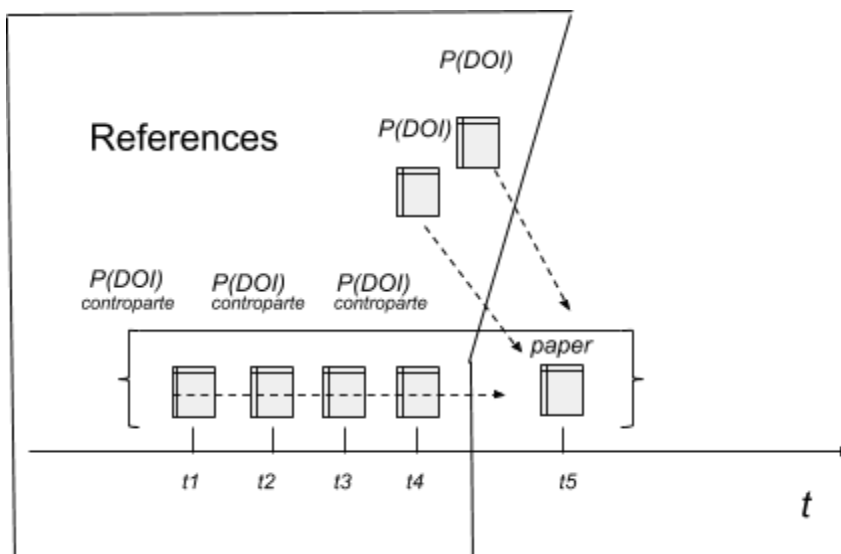
Se ogni controparte istanzia un DOI come proprietà, tale controparte diventa un oggetto pubblicato. Ciò che viene pubblicato, in questo caso, è una set di dati con relative meta-informazioni che ne caratterizzano il contenuto: il *data for paper*. Notate che questo prodotto è già pensato e finanziato per soddisfare questi criteri.

Se per *estensione di valore* intendiamo la capacità di ogni controparte di trasmettere le informazioni, la pubblicazione diretta di ogni controparte informativa è una soluzione possibile. Infatti, dal momento che nessuna delle controparti è funzionale ad una pubblicazione successiva, non siamo più soggetti al problema dell'ipotesi(A). In più, ogni controparte pubblicata può essere riusata da contesti che, al contrario, assumono la pubblicazione come direzione di produzione. Se questo è vero, pubblicare le controparti consente di estendere il valore delle attività produttive aumentando la possibilità di vedere diffuse le proprie attività attraverso pubblicazioni ottenute attraverso altri, esterni, processi di produzione



Può tale soluzione essere esportata in ambiti di ricerca publication oriented? a ben vedere sì. Il vantaggio di RFC* così descritto è che ogni controparte può essere riusata per la produzione di un articolo all'interno del processo di produzione originario. Il riuso, quindi, può essere una relazione interna. Come questa caratteristica del modello ci aiuta a superare il problema dell'ipotesi(A) in contesti publication oriented?

Come nel caso precedente, possiamo applicare un DOI ad ogni controparte, in modo tale che queste, durante il processo di produzione, possano essere pubblicate. In questo modo è possibile pubblicare un successivo articolo. In questo caso, è lo stesso processo produttivo che riusa una delle sua controparti temporali.



Se il problema è che non è possibile pubblicare dati che sono già stati pubblicati, il problema ha una soluzione in RFC*: infatti, i dati pubblicati in precedenza possono essere citati come ogni altro materiale citato all'interno. In più, poiché la controparte è sempre controparte nel processo produttivo, è lo stesso processo produttivo che cita uno dei suoi prodotti. Il criterio dell'originalità viene rispettato.

Questa soluzione presenta minori problemi implementativi rispetto alla precedente, poiché non richiede, ai domini publication oriented come quello accademico, di abbandonare la loro missione primaria. Al contrario il modello permette di perseguire la missione primaria con maggiore efficacia, estendendo alle proprie capacità produttive quel valore centrale, la pubblicazione, che quei domini richiedono.

6. Conclusione

In questo articolo abbiamo descritto e analizzato il modello Ready for Commons (RFC). Siamo partiti dalla descrizione del modello tradizionale di rappresentazione degli oggetti di conoscenza nella sezione 2. Abbiamo poi mostrato nella sezione 3. come tale modello non renda conto di alcuni aspetti legati ai processi di produzione della conoscenza. In particolare, di come non riesca a cogliere e sfruttare gli elementi di valore generati all'interno del processo, come nei casi di negatività. Abbiamo poi descritto il modello RFC di rappresentazione degli oggetti di conoscenza in 4 e come questo riesca, prima facie, ad includere aspetti lasciati fuori dal primo modello. Successivamente, in 4.3 abbiamo esposto alcuni limiti di RFC legati agli aspetti tradizionali dei modelli di produzione della conoscenza. Nel paragrafo 5 abbiamo proposto una possibile modifica del modello RFC, RFC*. Abbiamo esaminato due modi per risolvere i problemi descritti in 3. Sebbene il primo sia consistente, ha evidenti limiti di applicabilità. Il secondo, invece, ha più successo e ci consente di proporre una soluzione semplice ai problemi in 3.

7. Bibliografia

- Andow, D.A. (2003). Negative and positive data, statistical power, and confidence intervals. *Environ Biosafety Res.* 2003 Apr-Jun; 2(2): 75–80.
- Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90, 891–904.
- Heumann, K. F. (1956), Notes on negative data. *Amer. Doc.*, 7: 36–39

- Ney, A. (2014). *Metaphysics: An Introduction*. Routledge.

© Andrea Raimondi, Luca Corsato, Simone Cortesi 2018

DOI 10.5281/zenodo.1544011

Published by OSD opensorsdata srl | <https://osd.tools> | contact@osd.tools

This is an Open Access article distributed under the terms of the Creative Commons Attribution (<http://creativecommons.org/licenses/by/4.0/>)